# Realized Wavelet Jump-GARCH model: Can time-frequency decomposition of volatility improve its forecasting?☆

Jozef Barunik[a,b], Lukas Vacha[a,b]

[a] *Institute of Economic Studies, Charles University, Opletalova 21, 110 00, Prague, CR, tel. +420776259273*
[b] *Institute of Information Theory and Automation, Academy of Sciences of the Czech Republic, Pod Vodarenskou Vezi 4, 182 00, Prague, Czech Republic*



**Abstract**

In this paper, we propose a forecasting model for volatility based on its decomposition to several investment horizons and jumps. As a forecasting tool, we use Realized GARCH framework which models jointly returns and realized measures of volatility. Using jump wavelet two scale realized volatility estimator (JWTSRV), we first decompose the returns volatility into several investment horizons and jumps and then utilise this decomposition in a newly proposed Realized Jump-GARCH and Realized Wavelet-Jump GARCH models. On currency futures data covering the period of recent financial crisis we moreover compare the forecasts from Realized GARCH model using several additional realized volatility measures. Namely, we use the realized volatility, bipower variation, two-scale realized volatility, realized kernel and jump wavelet two scale realized volatility. We find that in-sample as well as out-of-sample performance of the model significantly differs based on the realized measure used. When JWTSRV estimator is used, model produces significantly best forecasts. Our Realized Wavelet-Jump GARCH model proves to further improve the volatility forecasts. We conclude that realized volatility measurement in the time-frequency domain and inclusion of jumps improves the volatility forecasting considerably.

*Keywords:* wavelet decomposition, jumps, volatility forecasting, Realized GARCH

*JEL: C14, C53, G17*

## 1. Introduction

Much of the recent popularity of realized volatility is mainly due to its two distinct implications for practical estimation and forecasting. The first relates to the measurement of realizations of the latent volatility process without the need for any assumptions about the explicit model. The second brings the possibility of forecasting volatility directly through standard time series econometrics with discretely sampled daily data, while effectively extracting information from intraday high-frequency data.

☆We are grateful to David Veredas and Karel Najzar for many useful comments and suggestions. We are also grateful to seminar participants at the Modeling High Frequency Data in Finance 3 in New York (July 2011) and Computational and Financial Econometrics in Oviedo (December 2012) for many useful discussions.

*Email address:* barunik@utia.cas.cz (Jozef Barunik )

The most fundamental result in realized variation states that it provides a consistent nonparametric estimate of price variability over a given time interval. The formalized theory is presented by Andersen et al. (2003). While these authors provide a unified framework for modeling, Zhou (1996) was one of the first to provide a formal assessment of the relationship between cumulative squared intraday returns and the underlying return variance. The pioneering work by Olsen & Associates on the use of high-frequency data, summarized by Dacorogna et al. (2001), produced milestone results for many of the more recent empirical developments in realized variation. A vast quantity of literature on several aspects of estimating volatility has emerged in the wake of these fundamental contributions. Our work builds on this popular *Realized Volatility* approach. While most time series models are set in the time domain, we enrich the analysis by the frequency domain. This is enabled by the use of the wavelet transform. It is a logical step to take, as the stock markets are believed to be driven by heterogeneous investment horizons. In our work, we ask if wavelet decomposition can improve our understanding of volatility series and hence improve volatility forecasting.

One very appealing feature of wavelets is that they can be embedded into stochastic processes, as shown by Antoniou and Gustafson (1999). Thus we can conveniently use them to extend the theory of realized measures as shown by Barunik and Vacha (2012). One of the common issues with the interpretation of wavelets in economic applications is that they are filter, thus they can hardly be used for forecasting in econometrics. Our wavelet-based estimator of realized volatility uses wavelets only to decompose the daily variation of the returns using intraday information, hence this is no longer an issue. As the wavelets are used to measure realized volatility at different investment horizons, our approach can be used to construct a forecasting model based on the wavelet decomposition.

Several attempts to use wavelets in the estimation of realized variation have emerged in the past few years. Hø g and Lunde (2003) were the first to suggest a wavelet estimator of realized variance. Capobianco (2004), for example, proposes to use a wavelet transform as a comparable estimator of quadratic variation. Subbotin (2008) uses wavelets to decompose volatility into a multi-horizon scale. Next, Nielsen and Frederiksen (2008) compare the finite sample properties of three integrated variance estimators, i.e., realized variance, Fourier and wavelet estimators. They consider several processes generating time series with a long memory, jump processes as well as bid-ask bounce. Gençay et al. (2010) mention the possible use of wavelet multiresolution analysis to decompose realized variance in their paper, while they concentrate on developing much more complicated structures of variance modeling in different regimes through wavelet-domain hidden Markov models. To complete the literature, Mancino and Sanfelici (2008), Olhede et al. (2009) propose estimators based on the Fourier transform. While the idea is very similar, this approach leads to realized volatility measurement in the frequency domain solely.

One exception which fully completes the current literature on using wavelets in realized variation theory is the work of Fan and Wang (2007), who were the first to use the wavelet-based realized variance estimator and also the methodology for the estimation of jumps from the data. In Barunik and Vacha (2012), we revisit and extend this work in a several ways. Instead of using the Discrete Wavelet Transform we use the Maximum Overlap Discrete Wavelet Transform, which is a more efficient estimator and is not restricted to sample sizes that are powers of two. We also use the Daubechies D(4) wavelet filter instead of the Haar type. Moreover, we bring large finite sample study confirming the behaviour of the wavelet estimators and run a forecasting simulation where our estimator confirms to improve forecasting of the integrated variance substantially when compared

to other estimators. Finally, in Barunik and Vacha (2012) we attempt to use the estimators to decompose stock market volatility into several investment horizons in a non-parametric way.

Motivated by these results, this paper focuses on proposing a model which will improve the forecasting of volatility. Similarly to Lanne (2007) and Andersen et al. (2011), we use the decomposition of the quadratic variation with the intention of building a more accurate forecasting model. Our approach is very different though, as we use wavelets to decompose the integrated volatility into several investment horizons and jumps. Moreover, we employ recently proposed realized GARCH framework of Hansen et al. (2011). Realized GARCH allows to model jointly returns and realized measures of volatility, while key feature is a measurement equation that relates the realized measure to the conditional variance of returns. We use several measures of realized volatility, namely realized volatility estimator proposed by Andersen et al. (2003), the bipower variation estimator of Barndorff-Nielsen and Shephard (2004), the two-scale realized volatility of Zhang et al. (2005), the realized kernel of Barndorff-Nielsen et al. (2008) and finally jump wavelet two-scale realized variance (JWTSRV) estimator of Barunik and Vacha (2012) in the framework of Realized GARCH and we find significant differences in volatility forecasts, while our JWTSRV estimator brings the largest improvement.

The main contribution of the paper are two new specifications of the Realized GARCH model based on the volatility decomposition and jumps. First, we utilize jumps estimated by the JWTSRV estimator to build a Realized Jump-GARCH(1,1) model. Second, we add a realized volatility measured at several investment horizons and build Realized Wavelet Jump-GARCH(1,1) models expecting that our models will result in better in-sample fits of the data as well as in out-of-sample forecasts. We are motivated by the statistical properties of the decomposed volatility series, which suggest that each scale might carry somewhat different information. The empirical analysis shows that our newly proposed models bring significant improvement in volatility forecasts.

The paper is organized in sections as follows. After the introduction, second Section reviews all the realized measures used in the forecasting excercise, third Section introduces our estimation of the realized variance and jumps using wavelets and Section four proposes a Realized Jump-GARCH(1,1) and Realized Wavelet Jump-GARCH(1,1) models. The fifth Section applies the presented theory, decomposes the empirical volatility of forex futures and finally uses the decomposition for forecasting.

## 2. Realized variance

We assume that the latent logarithmic asset price follows a standard jump-diffusion process and is contamined with microstructure noise. Let $y_t$ be the observed logarithmic prices at $0 \leq t \leq T$, which will be equal to the latent, so-called "true log-price process", $dp_t = \mu_t dt + \sigma_t dW_t + \xi_t dq_t$, and will contain microstructure noise,

$$y_t = p_t + \epsilon_t, \tag{1}$$

where $\epsilon_t$ is zero mean $i.i.d.$ noise with variance $\eta^2$, $q_t$ is a Poisson process uncorrelated with $W_t$ and governed by the constant jump intensity $\lambda$. The magnitude of the jump in the return process is controlled by factor $\xi_t \sim N(\bar{\xi}, \sigma_\xi^2)$.

Quadratic return variation over the $[t-h, t]$ time interval for $0 \leq h \leq t \leq T$, associated with

$p_t$,

$$QV_{t,h} = \underbrace{\int_{t-h}^{t} \sigma_s^2 ds}_{IV_{t,h}} + \underbrace{\sum_{t-h \le l \le t} J_l^2}_{JV_{t,h}} \tag{2}$$

can be naturally decomposed into two parts: integrated variance of the latent price process, $IV_{t,h}$ and jump variation $JV_{t,h}$. As detailed by Andersen et al. (2003), quadratic variation is a natural measure of variability in the logarithmic price.

A simple consistent estimator of the overall quadratic variation under the assumption of zero noise contamination in the price process is provided by the well-known realized variance, introduced by Andersen and Bollerslev (1998). The realized variance over $[t-h, t]$, for $0 \le h \le t \le T$, is defined by

$$\widehat{RV}_{t,h} = \sum_{i=1}^{N} r^2_{t-h+(\frac{i}{N})h}, \tag{3}$$

where $N$ is the number of observations in $[t-h, t]$ and $r_{t-h+(\frac{i}{N})h}$ is $i-$th intraday return in the $[t-h, t]$ interval. $\widehat{RV}_{t,h}$ converges in probability to $IV_{t,h} + JV_{t,h}$ as $N \to \infty$ (Andersen and Bollerslev, 1998; Andersen et al., 2001, 2003; Barndorff-Nielsen and Shephard, 2001, 2002a,b). As observed log-prices $y_t$ are contamined with noise in a real world and we are mainly interested in the $IV_{t,h}$ part of quadratic variation, subsequent literature has developed several estimators dealing with both jumps and noise.

### *2.1. Effect of microstructure noise*

Zhang et al. (2005) propose solution to the noise contamination by introducing the so-called two-scale realized volatility (TSRV henceforth) estimator. Authors propose a methodology for measurement of realized variance utilizing all of the available data using an idea of precise bias estimation. The two-scale realized variation over $[t-h, t]$, for $0 \le h \le t \le T$, is measured by

$$\widehat{RV}^{(TSRV)}_{t,h} = \underbrace{\widehat{RV}^{(average)}_{t,h}}_{\text{slow time scale}} - \frac{\bar{N}}{N} \underbrace{\widehat{RV}^{(all)}_{t,h}}_{\text{fast time scale}}, \tag{4}$$

where $\widehat{RV}^{(all)}_{t,h}$ is computed using Eq. (3) on all available data and $\widehat{RV}^{(average)}_{t,h}$ is constructed by averaging the estimators $\widehat{RV}^{(k)}_{t,h}$ obtained on $K$ grids of average size $\bar{N} = N/K$ as:

$$\widehat{RV}^{(average)}_{t,h} = \frac{1}{K} \sum_{k=1}^{K} \widehat{RV}^{(k)}_{t,h}. \tag{5}$$

In computing the TSRV, we have to first partition the original grid of observation times, $G = \{t_0, \ldots, t_N\}$, into subsamples $G^{(k)}$, $k = 1, \ldots, K$, where $N/K \to \infty$ as $N \to \infty$. For example, $G^{(1)}$ will start at the first observation and take an observation every 5 minutes, $G^{(2)}$ will start at the second observation and take an observation every 5 minutes, etc. Finally, we average these estimators through the subsamples, so we average the variation of the estimator as well. $\widehat{RV}^{(TSRV)}_{t,h}$ provides the first consistent and asymptotic estimator of the quadratic variation of $p_t$ with rate

of convergence $N^{-1/6}$. Zhang et al. (2005) also provide the theory for optimal choice of $K$ grids, $K^* = cN^{2/3}$, where the constant $c$ can be set to minimize the total asymptotic variance.

Another estimator, which is able to deal with the noise and which we use for the comparison in our study is the realized kernels (RK) estimator introduced by Barndorff-Nielsen et al. (2008). The realized kernel variance estimator is defined by

$$\widehat{RV}_{t,h}^{(RK)} = \gamma_{t,h,0} + \sum_{\eta=1}^{H} k\left(\frac{\eta-1}{H}\right)(\gamma_{t,h,\eta} + \gamma_{t,h,-\eta}), \tag{6}$$

with $\gamma_{t,h,\eta} = \sum_{i=1}^{N} r_{t-h+\left(\frac{i}{N}\right)h} r_{t-h+\left(\frac{i-\eta}{N}\right)h}$ denoting the $\eta$-th realized autocovariance with $\eta = -H, \ldots, -1, 0, 1, \ldots, H$ and $k(.)$ denotes the kernel function. Please note that for $\eta = 0$, $\gamma_{t,h,\eta} = \gamma_{t,h,0} = \widehat{RV}_{t,h}$ is estimate of the realized variance from Eq. (3). For the estimator to work, we need to choose the kernel function $k(.)$. In our study, we will focus on the Parzen kernel because it satisfies the smoothness conditions, $k'(0) = k'(1) = 0$ and is guaranteed to produce a non-negative estimate. The Parzen kernel function is given by

$$k(x) = \begin{cases} 1 - 6x^2 + 6x^3 & 0 \le x \le 1/2 \\ 2(1-x)^3 & 1/2 \le x \le 1. \\ 0 & x > 1 \end{cases} \tag{7}$$

We should note that the realized kernel estimator is computed without accounting for end effects, i.e. replacing the first and the last observation by local averages to eliminate the corresponding noise components (so-called "jittering"). Barndorff-Nielsen et al. (2008) argue that these effects are important theoretically, but are negligible practically.

*2.2. Effect of jumps*

By introducing the TSRV and the RK estimators, we will have benchmark estimators which are able to consistently estimate the quadratic variation from noisy observations. Still, we are interested to decompose quadratic variation into the integrated variance and jump variation component. Barndorff-Nielsen and Shephard (2004, 2006) develop a powerful and complete way of detecting the presence of jumps in high-frequency data. The basic idea is to compare two measures of the integrated variance, one containing the jump variation and the other being robust to jumps and hence containing only the integrated variation part. In our work, we use the Andersen et al. (2011) adjustment of the original Barndorff-Nielsen and Shephard (2004) estimator, which helps render it robust to certain types of microstructure noise. The bipower variation over $[t-h, t]$, for $0 \le h \le t \le T$, is defined by

$$\widehat{RV}_{t,h}^{(BV)} = \mu_1^{-2} \frac{N}{N-2} \sum_{i=3}^{N} |r_{t-h+\left(\frac{i-2}{N}\right)h}| . |r_{t-h+\left(\frac{i}{N}\right)h}|, \tag{8}$$

where $\mu_a = \pi/2 = E(|Z|^a)$, and $Z \sim N(0,1)$, $a \ge 0$ and $\widehat{RV}_{t,h}^{(BV)} \xrightarrow{p} \int_{t-h}^{t} \sigma_s^2 ds$. Thus $\widehat{RV}_{t,h}^{(BV)}$ provides a consistent estimator of the integrated variance and $\widehat{RV}_{t,h}^{(sparse)}$ provides a consistent estimator of the quadratic variation. Then, the jump variation can be estimated consistently as

the difference between the realized variance and the realized bipower variation:

$$\left(\widehat{RV}_{t,h}^{(sparse)} - \widehat{RV}_{t,h}^{(BV)}\right) \xrightarrow{p} \sum_{l=1}^{N_t} J_{t,h,l}^2. \tag{9}$$

Under the assumption of no jump and some other regularity conditions, Barndorff-Nielsen and Shephard (2006) provided the joint asymptotic distribution of the jump variation. Under the null hypothesis of no within-day jumps,

$$Z_{t,h} = \frac{\frac{\widehat{RV}_{t,h}^{(sparse)} - \widehat{RV}_{t,h}^{(BV)}}{\widehat{RV}_{t,h}^{(sparse)}}}{\sqrt{((\frac{\pi}{2})^2 + \pi - 5)\frac{1}{n}\max\left(1, \frac{\widehat{TQ}_{t,h}}{\left(\widehat{RV}_{t,h}^{(BV)}\right)^2}\right)}}, \tag{10}$$

where $\widehat{TQ}_{t,h} = N\mu_{4/3}^{-3}(\frac{N}{n-4})\sum_{j=5}^{N}|r_{t-h+(\frac{i-4}{N})h}|^{4/3}|r_{t-h+(\frac{i-2}{N})h}|^{4/3}|r_{t-h+(\frac{i-2}{N})h}|^{4/3}$ is asymptotically standard normally distributed. Using this theory, the contribution of the jump variation to the quadratic variation of the price process is measured by

$$\widehat{J}_{t,h} = I_{Z_{t,h}>\Phi_\alpha}\left(\widehat{RV}_{t,h}^{(sparse)} - \widehat{RV}_{t,h}^{(BV)}\right), \tag{11}$$

where $I_{Z_{t,h}>\Phi_\alpha}$ denotes the indicator function and $\Phi_\alpha$ refers to the chosen critical value from the standard normal distribution. The measure of integrated variance is defined as

$$\widehat{C}_{t,h} = I_{Z_{t,h}\leq\Phi_\alpha}\widehat{RV}_{t,h}^{(sparse)} + I_{Z_{t,h}>\Phi_\alpha}\widehat{RV}_{t,h}^{(BV)}, \tag{12}$$

ensuring that the jump measure and the continuous part add up to the estimated variance without jumps.

We use the described jump detection methodology as the benchmark and we focus on wavelet methods for detecting jumps in the data, as described in the following section.

## 3. Estimation of the realized variance using wavelets

While most time series models are naturally set in the time domain, wavelet transform help us to enrich the analysis of realized variance by the frequency domain. It is a logical step to take, as the stock markets are believed to be driven by heterogeneous investment horizons, so volatility dynamics should be understood not only in time but at investment horizons as well. We will introduce general ideas of constructing the estimators here, while we keep the details necessary to understand the wavelet theory in the Appendix Appendix A.

Following Barunik and Vacha (2012), the wavelet-based realized variance over $[t-h, t]$, for $0 \leq h \leq t \leq T$, is defined by

$$\widehat{RV}_{t,h}^{(WRV)} = \sum_{j=1}^{J^m+1}\sum_{k=1}^{N}\mathcal{W}_{j,t-h+\frac{k}{N}h}^2, \tag{13}$$

where $N$ is the number of intraday observations in $[t-h,t]$ and $J^m$ is the number of scales we consider. $\mathcal{W}_{j,t-h+\frac{k}{N}h}$ are the MODWT coefficients, unaffected by boundary conditions, defined in Eq.(A.7) on returns data $r_{t,h}$ on components $j=1,\ldots,J^m+1$, where $J^m \leq \log_2 N$. With increasing sampling frequency $N\to\infty$ the wavelet-based realized variance estimator is unbiased and consistent estimator of the quadratic variation, and it is the same as realized variance estimator; for more details see Barunik and Vacha (2012). Still, we assume that observed $y_t$ process contains noise as well as jumps. Therefore, we need to introduce the concepts which will be able to deal with both.

### *3.1. Realized jump estimation using wavelets*

Wavelets can also be used for estimating jumps and separating integrated variance from jump variation. We assume that the sample path of $p_t$ has a finite number of jumps (*a.s.*). Following the theoretical results of Wang (1995) on the wavelet jump detection of the deterministic functions with *i.i.d.* additive noise $\epsilon_t$, we use MODWT as the discretized version of the continuous wavelet transform. Unlike the ordinary DWT, the MODWT is not restricted to a dyadic sample length. For the estimation of jump location we use the universal threshold (Donoho and Johnstone, 1994) on the first level wavelet coefficients of $y_t$ over $[t-h,t]$, $\mathcal{W}_{1,k}$. If for some $\mathcal{W}_{1,k}$

$$|\mathcal{W}_{1,k}| > d\sqrt{2\log N}, \tag{14}$$

then $\hat{\tau}_l = \{k\}$ is the estimated jump location with size $\bar{y}_{\hat{\tau}_{l+}} - \bar{y}_{\hat{\tau}_{l-}}$ (averages over $[\hat{\tau}_l, \hat{\tau}_l+\delta_n]$ and $[\hat{\tau}_l, \hat{\tau}_l-\delta_n]$, respectively, with $\delta_n>0$ being the small neighborhood of the estimated jump location[1] $\hat{\tau}_l \pm \delta_n$) and where $d$ is median absolute deviation estimator defined as $(2^{1/2})median\{|\mathcal{W}_{1,k}|, k=1,\ldots,N\}/0.6745$ (Percival and Walden, 2000).

Using the result of Fan and Wang (2007), the jump variation is then estimated by the sum of the squares of all the estimated jump sizes:

$$\widehat{JV}^W_{t,h} = \sum_{l=1}^{N_t} (\bar{y}_{t,h,\hat{\tau}_{l+}} - \bar{y}_{t,h,\hat{\tau}_{l-}})^2, \tag{15}$$

thus we are able to estimate the jump variation from the process consistently with the convergence rate $N^{-1/4}$. In the following analysis, we will be able to separate the continuous part of the price process containing noise from the jump variation. This result can be found in Fan and Wang (2007) and it states that the jump-adjusted process $y^{(J)}_{t,h} = y_{t,h} - \widehat{JV}^W_{t,h}$ converges in probability to the continuous part without jumps. Thus, if we are able to deal with the noise in $y^{(J)}_{t,h}$, we will be able to estimate the $IV_{t,h}$.

### *3.2. Jump wavelet two scale realized variance (JWTSRV) estimator*

In the final estimator, we utilize the TSRV estimator of Zhang et al. (2005), the wavelet realized variance estimator Eq. (13) and the wavelet jump detection method. Final estimator will moreover decompose the integrated variance into $J^m+1$ components so we will be able to study

[1] Due to the nature of the MODWT filters, we need to correct the position of the wavelet coefficient to get the precise position of the jump. For more details see Percival and Mofjeld (1997).

the dynamics of volatility at various investment horizons.

Following Barunik and Vacha (2012), we define the jump-adjusted wavelet two-scale realized variance (JWTSRV) estimator over $[t-h,t]$, for $0 \leq h \leq t \leq T$, on the observed jump-adjusted data, $y_{t,h}^{(J)} = y_{t,h} - \sum_{l=1}^{N_t} J_l$ as:

$$\widehat{RV}_{t,h}^{(JWTSRV)} = \sum_{j=1}^{J^m+1} \widehat{RV}_{j,t,h}^{(JWTSRV)} = \sum_{j=1}^{J^m+1} \left( \widehat{RV}_{j,t,h}^{(W,J)} - \frac{\bar{N}}{N} \widehat{RV}_{j,t,h}^{(WRV,J)} \right), \tag{16}$$

where $\widehat{RV}_{j,t,h}^{(W,J)} = \frac{1}{G} \sum_{g=1}^{G} \sum_{k=1}^{N} \mathcal{W}_{j,t-h+\frac{k}{N}h}^2$ is obtained from wavelet coefficient estimates on a grid of size $\bar{N} = N/G$ and $\widehat{RV}_{j,t,h}^{(WRV,J)} = \sum_{k=1}^{N} \mathcal{W}_{j,t-h+\frac{k}{N}h}^2$ is the wavelet realized variance estimator at a scale $j$ on the jump-adjusted observed data, $y_{t,h}^{(J)}$.

The JWTSRV estimator decomposes the realized variance into an arbitrary chosen number of investment horizons and jumps. Barunik and Vacha (2012) discuss that it is consistent estimator of the integrated variance as it converges in probability to the integrated variance of the process $p_t$. Barunik and Vacha (2012) also test the small sample performance of the estimator in a large Monte Carlo study and they find that it is able to recover true integrated variance from the noisy process with jumps very precisely. They also run a forecasting simulation where JWTSRV estimator confirms to improve forecasting of the integrated variance substantially. In small samples, a small sample refinement can be constructed (Zhang et al., 2005):

$$\widehat{RV}_{t,h}^{(JWTSRV,adj)} = \left(1 - \frac{\bar{N}}{N}\right)^{(-1)} \widehat{RV}_{t,h}^{(JWTSRV)}. \tag{17}$$

When referring to the realized volatility estimated using our JWTSRV estimator, we will refer to the $\sqrt{\widehat{RV}_{t,h}^{(JWTSRV,adj)}}$.

## 4. A forecasting model based on decomposed integrated volatilities

Similarly to Lanne (2007) and Andersen et al. (2011), we use the decomposition of the quadratic variation with the intention of building a more accurate forecasting model. Our approach is very different though, as we use wavelets to decompose the integrated volatility into several investment horizons and jumps first. Moreover, we employ recently proposed Realized GARCH framework of Hansen et al. (2011). Realized GARCH allows to model jointly returns and realized measures of volatility, while key feature is a measurement equation that relates the realized measure to the conditional variance of returns. We expect that our modification will result in better in-sample fits of the data as well as out-of-sample forecasts.

### *4.1. Realized GARCH framework for forecasting*

The key object of interest in GARCH family is the conditional variance, $h_t = var(r_t|\mathcal{F}_{t-1})$, where $r_t$ is a time series of returns. While in a standard GARCH(1,1) model the conditional variance, $h_t$ is dependent on its past $h_{t-1}$ and $r_{t-1}^2$, Hansen et al. (2011) propose to utilize realized

measures of volatility and make $h_t$ dependent on them as well. Authors propose so-called measurement equation which ties the realized measure to latent volatility. The general framework of Realized GARCH$(p, q)$ models is well connected to existing literature in Hansen et al. (2011). Here, we restrict ourselves to the simple log-linear specification of Realized GARCH$(1, 1)$ with Gaussian innovations which we will use to build our model. A simple log-linear Realized GARCH$(1, 1)$ model is given by

$$\begin{aligned} r_t &= \sqrt{h_t} z_t, && (18) \\ \log(h_t) &= \omega + \beta \log(h_{t-1}) + \gamma \log(x_{t-1}) && (19) \\ \log(x_t) &= \xi + \psi \log(h_t) + \tau_1 z_t + \tau_2 z_t^2 + u_t, && (20) \end{aligned}$$

where $r_t$ is the return, $x_t$ a realized measure of volatility, $z_t \sim i.i.d(0, 1)$ and $u_t \sim i.i.d(0, \sigma_u^2)$ with $z_t$ and $u_t$ being mutually independent, $h_t = var(r_t|\mathcal{F}_{t-1})$ with $\mathcal{F}_t = \sigma(r_t, x_t, r_{t-1}, x_{t-1}, \ldots)$ and $\tau(z) = \tau_1 z_t + \tau_2 z_t^2$ is called leverage function.

It is worth noting that while we use only this specific version of Realized GARCH, Hansen et al. (2011) introduces a general family of models which generalized a GARCH models as it can nest any GARCH specification. Also assumption on innovations is not essential and can be changed to other common assumptions as Student's $t$ for example.

Hansen et al. (2011) provide also the asymptotic properties of the quasi-maximum likelihood estimator (QMLE henceforth) and propose to use it for the parameter estimation. The structure of the QMLE is very similar to that of the standard GARCH model, but we need to accommodate also realized measures in the estimation. The log-likelihood function is given by

$$\log L(\{r_t, x_t\}_{t=1}^T; \theta) = \sum_{t=1}^T \log f(r_t, x_t|\mathcal{F}_{t-1}). \tag{21}$$

Standard GARCH models do not have realized measure $x_t$, so we need to factorize the joint conditional density

$$f(r_t, x_t|\mathcal{F}_{t-1}) = f(r_t|\mathcal{F}_{t-1}) f(x_t|r_t, \mathcal{F}_{t-1}). \tag{22}$$

and use the partial log-likelihood, $\ell(r) = \sum_{t=1}^T \log f(r_t|\mathcal{F}_{t-1})$ when comparing the fits to a standard GARCH. For the Gaussian specification of $z_t$ and $u_t$, the joint likelihood is then split into the sum

$$\ell(r, x) = \underbrace{-0.5 \sum_{t=1}^T \left(\log(2\pi) + \log(h_t) + r_t^2/h_t\right)}_{\ell(r)} + \underbrace{-0.5 \sum_{t=1}^T \left(\log(2\pi) + \log(\sigma_u^2) + u_t^2/\sigma_u^2\right)}_{\ell(x|r)} \tag{23}$$

while other standard specifications can be used as well in a similar manner.

Realized GARCH framework is rather general. For example, it allows to accommodate more realized measures. In our analysis, we will estimate Realized GARCH(1,1) models using different $x_t$, namely RV, BV, TSRV, RK and JWTSRV from previous sections and compare its performance.

### *4.2. Realized Wavelet Jump-GARCH(1,1)*

By estimating different Realized GARCH models using various realized measures, we will see which measure carries the best information for forecasting of volatility. In addition, we would like

to utilize estimated jumps as well as decomposition of JWTSRV and propose two more specifications, Realized Jump-GARCH(1,1) model (Realized J-GARCH) and Realized Wavelet Jump-GARCH(1,1) model (Realized WJ-GARCH). First, by addition of estimated jumps into the variance equation, we obtain Realized J-GARCH(1,1) model given by

$$\begin{aligned} r_t &= \sqrt{h_t} z_t, && (24)\\ \log(h_t) &= \omega + \beta \log(h_{t-1}) + \gamma \log(x_{t-1}) + \gamma_J \log(1 + JV_{t-1}), && (25)\\ \log(x_t) &= \xi + \psi \log(h_t) + \tau_1 z_t + \tau_2 z_t^2 + u_t, && (26) \end{aligned}$$

where $x_t$ and $JV_t$ is estimated using Eq. (16) and Eq. (15) by our $\widehat{RV}_t^{(JWTSRV)}$ and $\widehat{JV}_t^W$ respectively and $z_t$ and $u_t$ come from Gaussian normal distribution and are mutually independent. This model is logical step in generalizing the Realized GARCH structure as $\widehat{RV}_t^{(JWTSRV)}$ and $\widehat{JV}_t^W$ add up to a quadratic variation of underlying price process which is not biased by noise. If jumps have a significant impact on volatility forecasts, $\gamma_J$ coefficient should be significantly different from zero.

Finally, we utilize a wavelet decomposition of integrated volatility to different investment horizons and estimate the model where $x_{j,t}$ will represent $\widehat{RV}_{j,t}^{(JWTSRV)}$ at all estimated investment horizons $j = 1, ..., J^m + 1$. The Realized Wavelet Jump-GARCH(1,1) model is given by

$$\begin{aligned} r_t &= \sqrt{h_t} z_t, && (27)\\ \log(h_t) &= \omega + \beta \log(h_{t-1}) + \sum_{j=1}^{J^m+1} \gamma_{W_j} \log(x_{j,t-1}) + \gamma_J \log(1 + JV_{t-1}), && (28)\\ \log(x_t) &= \xi + \psi \log(h_t) + \tau_1 z_t + \tau_2 z_t^2 + u_t, && (29) \end{aligned}$$

where $x_{j,t}$ is estimated using Eq. (16) by $\widehat{RV}_{j,t}^{(JWTSRV)}$, $JV_t$ is estimated using Eq. (15) by $\widehat{JV}_t^W$ and $z_t$ and $u_t$ come from Gaussian normal distribution and are mutually independent. Note that $j$ components always add up to overall variance $x_t = \widehat{RV}_t^{(JWTSRV)}$. Our last model is motivated by the decomposition of realized volatility into several investment horizons. $\gamma_{W_j}$ will provide a good guide for significance of various investment horizons.

All the models are estimated by QMLE and can be easily generalized by assuming different distributions of $z_t$ and $u_t$. We have also tried to incorporate different distributions[2] but the results did not change qualitatively and to keep the number of estimated models under control, we report the results for the Gaussian case only.

*4.3. Forecast evaluation using different realized variance measures*

To analyze the forecast efficiency and information content of different volatility estimators in the Realized GARCH framework, we employ the popular approach of Mincer and Zarnowitz (1969) regressions. The regression takes the form:

$$V_{t+1}^{(m)} = \alpha + \beta V_t^{RG-(k)} + \epsilon_t, \quad (30)$$

[2]Results for other cases are available from authors upon request.

Table 1: The table summarizes the daily log-return distributions of GBP, CHF and EUR futures. The sample period extends from January 5, 2007 through November 17, 2010, accounting for a total of 944 observations.

| | Mean | St.dev. | Skew. | Kurt. |
|---|---|---|---|---|
| **GBP** | 0.0001 | 0.0119 | -0.3852 | 4.4356 |
| **CHF** | 0.0002 | 0.0068 | 0.2440 | 5.4662 |
| **EUR** | 0.0002 | 0.0099 | 0.1536 | 4.4951 |

with $V_{t+1}^{(m)}$ being the integrated volatility estimated using the square root of the $m$-th estimator, namely, RV, BV, TSRV, RK and JWTSRV, respectively. $V_t^{RG-(k)}$ denotes the 1-day ahead forecast of $V_{t+1}^{(m)}$ using the $k$-th estimator based on Realized GARCH(1,1), namely RV, BV, TSRV, RK, JWTSRV and finally Realized J-GARCH(1,1) and Realized WJ-GARCH(1,1). We report in-sample as well as rolling out-of-sample results.

After testing the forecasting efficiency of the different volatility models we would also like to test the information content of the wavelet decomposition of the realized volatility. For this purpose, we separately estimate Realized J-GARCH$(1,1)$ for all components $JWTSRV_j$ for $j = 1, \ldots, 5$ of the realized volatility. Finally, we use Heteroskedasticity-adjusted Mean Square Error (HMSE) of Bollerslev and Ghysels (1996) and QLIKE of Bollerslev et al. (1994).

## 5. Does decomposition bring any improvement in volatility forecasting?

### *5.1. Data description*

Foreign exchange future contracts are traded on the Chicago Mercantile Exchange (CME) on a 24-hour basis. As these markets are among the most liquid, they are suitable for analysis of high-frequency data. We will estimate the realized volatility of British pound (GBP), Swiss franc (CHF) and euro (EUR) futures. All contracts are quoted in the unit value of the foreign currency in US dollars. It is advantageous to use currency futures data for the analysis instead of spot currency prices, as they embed interest rate differentials and do not suffer from additional microstructure noise coming from over-the-counter trading. The cleaned data are available from Tick Data, Inc.[3]

It is very important to look first at the changes in the trading system before we proceed with the estimation on the data. In August 2003, for example, the CME launched the Globex trading platform, and for the first time ever in a single month, the trading volume on the electronic trading platform exceeded 1 million contracts every day. On Monday, December 18, 2006, the CME Globex(R) electronic trading platform started offering nearly continuous trading. More precisely, the trading cycle became 23 hours a day (from 5:00 pm on the previous day until 4:00 pm on current day, with a one-hour break in continuous trading), from 5:00 pm on Sunday until 4:00 pm on Friday. These changes certainly had a dramatic impact on trading activity and the amount of information available, resulting in difficulties in comparing the estimators on the pre-2003 data, the 2003–2006 data and the post–2006 data. For this reason, we restrict our analysis to a sample period extending from January 5, 2007 through November 17, 2010, which contains the most recent financial crisis. The futures contracts we use are automatically rolled over to provide continuous price records, so we do not have to deal with different maturities.

[3] http://www.tickdata.com/

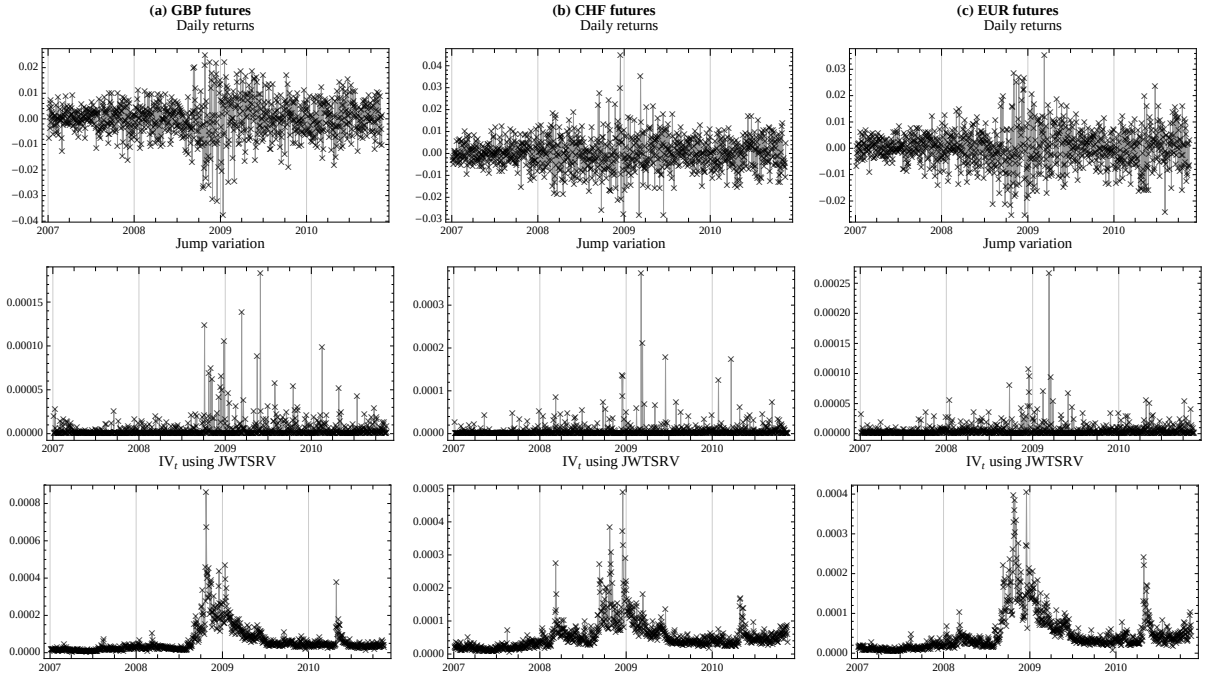


Figure 1: Daily returns, estimated jump variation and $IV_t$ estimated by JWTSRV for (a) GBP, (b) CHF and (c) EUR futures.

The tick-by-tick transactions are recorded in Chicago Time, referred to as Central Standard Time (CST). Therefore, in a given day, trading activity starts at 5:00 pm CST in Asia, continues in Europe followed by North America, and finally closes at 4:00 pm in Australia. To exclude potential jumps due to the one-hour gap in trading, we redefine the day in accordance with the electronic trading system. Moreover, we eliminate transactions executed on Saturdays and Sundays, US federal holidays, December 24 to 26, and December 31 to January 2, because of the low activity on these days, which could lead to estimation bias. Finally, we are left with 944 days in the sample. Looking more deeply at higher frequencies, we find a large amount of multiple transactions happening exactly at the same time stamp. We use the arithmetic average for all observations with the same time stamp. Table 1 presents the summary statistics for the daily log-returns of GBP, CHF and EUR futures over the sample period, $t = 1, \ldots, 944$, i.e., January 5, 2007 to November 17, 2010. The summary statistics display an average return very close to zero, skewness, and excess kurtosis which is consistent with the large empirical literature.

Having prepared the data, we can estimate the integrated volatility using different estimators and use them within proposed forecasting framework. For each futures contract, the daily integrated volatility is estimated using the square root of realized variance estimator of Andersen et al. (2003), the bipower variation estimator of Barndorff-Nielsen and Shephard (2004), the two-scale realized volatility of Zhang et al. (2005), the realized kernel of Barndorff-Nielsen et al. (2008) described in the Section 2. Finally, we utilize our jump wavelet two-scale realized variance estimator defined by Eq. (16). All the estimators are adjusted for small sample bias. For convenience, we refer to the estimators in the description of the results as RV, BV, TSRV, RK and JWTSRV, respectively. The RV and BV estimates are estimated on 5-min log-returns. The TSRV and the

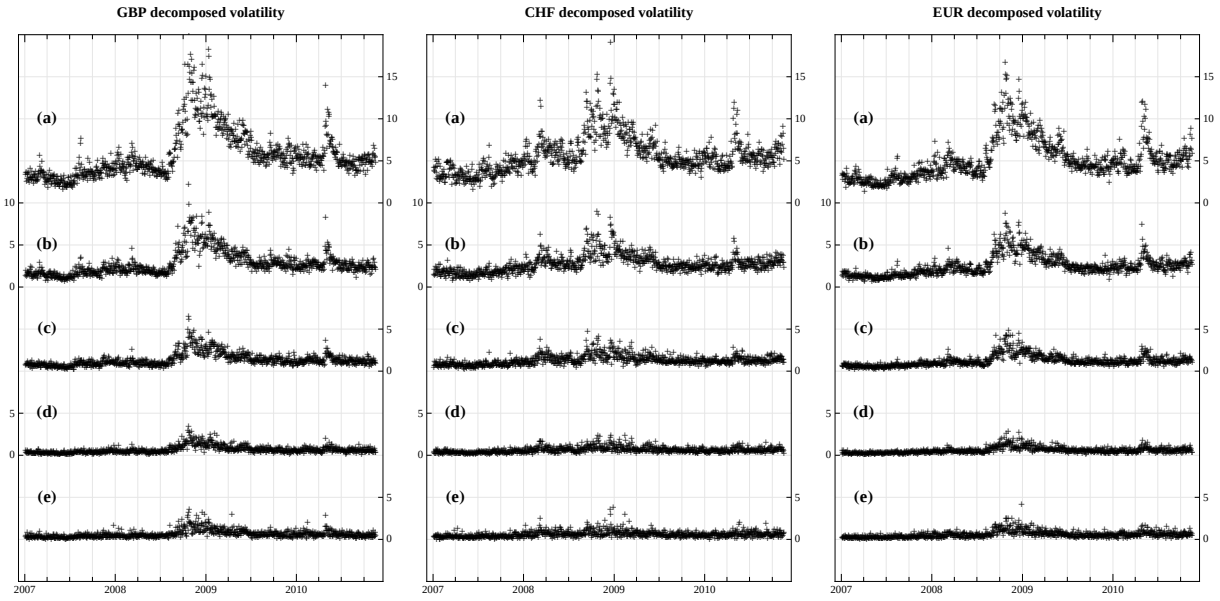


Figure 2: Decomposed annualized volatility (by 252 days) of GBP, CHF and EUR futures using JWTSRV, (a) volatility on investment horizon of 10 minutes, (b) volatility on investment horizon of 20 minutes, (c) volatility on investment horizon of 40 minutes, (d) volatility on investment horizon of 80 minutes, (e) volatility on investment horizon up to 1 day. Note that sum of components (a), (b), (c), (d) and (e) give total volatility.

JWTSRV are estimated using a slow time scale of 5 minutes.

The decomposition of volatility into the so-called continuous and jump part is depicted by Figure 1, which provide the returns, estimated jumps and finally integrated variances using JWTSRV estimator for all three futures pairs. Figure 2 shows the further decomposition into several investment horizons. For better illustration, we annualize the square root of the integrated variance in order to get the annualized volatility and we compute the components of the volatility on several investment horizons. Figure 2 (a) to (e) show the investment horizons of 10 minutes, 20 minutes, 40 minutes, 80 minutes and up to 1 day, respectively. It is very interesting that most of the volatility (around 50%) comes from the fast, 10-minute investment horizon which is a new insight. In fact, it is a logical finding, as it shows that volatility is created on fast scales of up to 10 minutes rather than on slower scales. The longer the horizon, the lower the contribution of the variance to the total variation.

### *5.2. Forecasting results*

We present the main results of estimation and forecasting here. The estimation strategy is as follows. For each of three forex futures considered, namely GBP, CHF and EUR, we first estimate benchmark GARCH(1,1) model. Then, we estimate the Realized GARCH of Hansen et al. (2011). It is important to note that we would like to compare performance of the model with several realized volatility measures, namely RV, BV, RK, TSRV and JWTSRV. Finally, we add our Realized Jump-GARCH model and Realized Wavelet Jump-GARCH model. We use the period from January 5, 2007 to February 2, 2010 for estimation of all the models. Thus, we refer to this period as the in-sample period. The rest of the year 2010 is saved for comparison of the out-of-sample forecasts on a rolling basis. We use open-to-close returns as well as open-to-close realized measures in the analysis.

Table 2: Results for the GBP futures: in-sample fits of GARCH(1,1), Realized GARCH(1,1) with RV, BV, RK, TSRV, JWTSRV, Realized (Wavelet)Jump-GARCH denoted as RJ-G and RWJ-G and Realized Jump-GARCH on $JWTSRV_j$ decompositions. Robust standard errors are reported in parentheses.

| | GARCH | Realized GARCH | | | | | Realized (W)J-GARCH | | Realized Jump-GARCH on $JWTSRV_j$ | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| In-sample | | RV | BV | RK | TSRV | JWTSRV | RJ-G | RWJ-G | $j=1$ | $j=2$ | $j=3$ | $j=4$ | $j=5$ |
| $\omega$ | 0.007 | 0.147 | 0.155 | 0.130 | 0.153 | 0.195 | 0.194 | 0.194 | 0.174 | 0.165 | 0.156 | 0.113 | 0.095 |
| | (0.004) | (0.025) | (0.029) | (0.024) | (0.030) | (0.035) | (0.033) | (0.027) | (0.039) | (0.032) | (0.030) | (0.021) | (0.019) |
| $\alpha$ | 0.052 | | | | | | | | | | | | |
| | (0.014) | | | | | | | | | | | | |
| $\beta$ | 0.941 | 0.599 | 0.568 | 0.691 | 0.573 | 0.494 | 0.496 | 0.511 | 0.524 | 0.562 | 0.626 | 0.751 | 0.806 |
| | (0.016) | (0.037) | (0.039) | (0.035) | (0.041) | (0.037) | (0.041) | (0.042) | (0.044) | (0.041) | (0.041) | (0.028) | (0.026) |
| $\gamma$ | | 0.271 | 0.283 | 0.207 | 0.288 | 0.339 | 0.351 | | | | | | |
| | | (0.033) | (0.033) | (0.028) | (0.035) | (0.033) | (0.036) | | | | | | |
| $\gamma_J$ | | | | | | | 0.043 | 0.036 | 0.049 | 0.075 | 0.092 | 0.108 | 0.047 |
| | | | | | | | (0.059) | (0.063) | (0.069) | (0.080) | (0.074) | (0.072) | (0.086) |
| $\gamma_{W_1}$ | | | | | | | | 0.249 | 0.331 | | | | |
| | | | | | | | | (0.049) | (0.045) | | | | |
| $\gamma_{W_2}$ | | | | | | | | 0.006 | | 0.298 | | | |
| | | | | | | | | (0.043) | | (0.036) | | | |
| $\gamma_{W_3}$ | | | | | | | | 0.064 | | | 0.252 | | |
| | | | | | | | | (0.028) | | | (0.033) | | |
| $\gamma_{W_4}$ | | | | | | | | 0.007 | | | | 0.121 | |
| | | | | | | | | (0.021) | | | | (0.022) | |
| $\gamma_{W_5}$ | | | | | | | | 0.012 | | | | | 0.121 |
| | | | | | | | | (0.017) | | | | | (0.019) |
| $\xi$ | | -0.542 | -0.549 | -0.626 | -0.533 | -0.577 | -0.560 | -0.534 | -0.536 | -0.567 | -0.642 | -0.744 | -0.811 |
| | | (0.056) | (0.073) | (0.074) | (0.074) | (0.064) | (0.065) | (0.066) | (0.065) | (0.072) | (0.070) | (0.077) | (0.076) |
| $\psi$ | | 1.439 | 1.490 | 1.456 | 1.445 | 1.456 | 1.398 | 1.392 | 1.399 | 1.418 | 1.426 | 1.483 | 1.538 |
| | | (0.105) | (0.108) | (0.100) | (0.105) | (0.119) | (0.094) | (0.088) | (0.112) | (0.102) | (0.101) | (0.111) | (0.105) |
| $\sigma_u$ | | 0.313 | 0.308 | 0.379 | 0.318 | 0.293 | 0.291 | 0.290 | 0.292 | 0.328 | 0.389 | 0.469 | 0.328 |
| | | (0.080) | (0.080) | (0.080) | (0.080) | (0.080) | (0.080) | (0.080) | (0.080) | (0.080) | (0.080) | (0.080) | (0.080) |
| $\tau_1$ | | -0.041 | -0.046 | -0.030 | -0.043 | -0.046 | -0.046 | -0.054 | -0.059 | -0.043 | -0.038 | -0.029 | -0.019 |
| | | (0.012) | (0.012) | (0.014) | (0.013) | (0.011) | (0.011) | (0.011) | (0.011) | (0.013) | (0.015) | (0.017) | (0.022) |
| $\tau_2$ | | 0.077 | 0.064 | 0.107 | 0.080 | 0.063 | 0.066 | 0.056 | 0.056 | 0.059 | 0.078 | 0.083 | 0.150 |
| | | (0.009) | (0.009) | (0.011) | (0.009) | (0.008) | (0.008) | (0.008) | (0.008) | (0.009) | (0.011) | (0.013) | (0.016) |
| $l(r,x)$ | | -1185.123 | -1172.427 | -1327.555 | -1196.632 | -1134.104 | -1122.474 | -1120.224 | -1126.749 | -1214.461 | -1339.218 | -1478.777 | -1605.582 |
| $l(r)$ | -1001.735 | -992.267 | -992.605 | -992.389 | -992.828 | -990.871 | -984.540 | -984.410 | -985.216 | -986.411 | -983.982 | -983.646 | -986.431 |
| out-of-sample | | | | | | | | | | | | | |
| $\alpha_{MZ}$ | -0.673 | -0.251 | -0.325 | -0.215 | -0.261 | -0.313 | -0.263 | -0.266 | -0.269 | -0.294 | -0.263 | -0.326 | -0.275 |
| $\beta_{MZ}$ | 1.607 | 0.946 | 1.026 | 0.874 | 0.963 | 0.994 | 0.952 | 0.959 | 0.966 | 0.983 | 0.965 | 1.026 | 0.983 |
| $R^2_{MZ}$ | 0.083 | 0.339 | 0.399 | 0.262 | 0.352 | 0.422 | 0.426 | 0.422 | 0.426 | 0.416 | 0.359 | 0.326 | 0.336 |
| HMSE | 2.877 | 1.674 | 1.758 | 2.307 | 1.655 | 1.744 | 1.669 | 1.747 | 1.717 | 1.781 | 2.448 | 2.448 | 2.477 |
| QLIKE | 0.488 | 0.452 | 0.471 | 0.432 | 0.463 | 0.413 | 0.400 | 0.400 | 0.397 | 0.408 | 0.423 | 0.435 | 0.422 |

Table 3: Results for the CHF futures: in-sample fits of GARCH(1,1), Realized GARCH(1,1) with RV, BV, RK, TSRV, JWTSRV, Realized (Wavelet)Jump-GARCH denoted as RJ-G and RWJ-G and Realized Jump-GARCH on $JWTSRV_j$ decompositions. Robust standard errors are reported in parentheses.

| | GARCH | Realized GARCH | | | | | Realized (W)J-GARCH | | Realized Jump-GARCH on $JWTSRV_j$ | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| In-sample | | RV | BV | RK | TSRV | JWTSRV | RJ-G | RWJ-G | $j=1$ | $j=2$ | $j=3$ | $j=4$ | $j=5$ |
| $\omega$ | 0.004 | 0.034 | 0.038 | 0.054 | 0.022 | 0.014 | 0.006 | 0.024 | 0.001 | 0.013 | 0.016 | 0.039 | 0.065 |
| | (0.003) | (0.014) | (0.018) | (0.012) | (0.015) | (0.037) | (0.022) | (0.027) | (0.018) | (0.017) | (0.015) | (0.013) | (0.015) |
| $\alpha$ | 0.060 | | | | | | | | | | | | |
| | (0.011) | | | | | | | | | | | | |
| $\beta$ | 0.939 | 0.711 | 0.680 | 0.771 | 0.710 | 0.590 | 0.593 | 0.588 | 0.603 | 0.669 | 0.747 | 0.782 | 0.820 |
| | (0.010) | (0.033) | (0.034) | (0.028) | (0.032) | (0.023) | (0.042) | (0.044) | (0.041) | (0.041) | (0.029) | (0.029) | (0.027) |
| $\gamma$ | | 0.302 | 0.323 | 0.234 | 0.297 | 0.421 | 0.426 | | | | | | |
| | | (0.037) | (0.039) | (0.028) | (0.034) | (0.063) | (0.049) | | | | | | |
| $\gamma_J$ | | | | | | | 0.063 | 0.046 | 0.069 | 0.074 | 0.102 | 0.086 | 0.084 |
| | | | | | | | (0.052) | (0.054) | (0.052) | (0.053) | (0.056) | (0.049) | (0.065) |
| $\gamma_{W_1}$ | | | | | | | | 0.218 | 0.407 | | | | |
| | | | | | | | | (0.051) | (0.050) | | | | |
| $\gamma_{W_2}$ | | | | | | | | 0.114 | | 0.342 | | | |
| | | | | | | | | (0.057) | | (0.046) | | | |
| $\gamma_{W_3}$ | | | | | | | | 0.054 | | | 0.261 | | |
| | | | | | | | | (0.040) | | | (0.033) | | |
| $\gamma_{W_4}$ | | | | | | | | 0.014 | | | | 0.182 | |
| | | | | | | | | (0.027) | | | | (0.028) | |
| $\gamma_{W_5}$ | | | | | | | | 0.026 | | | | | 0.182 |
| | | | | | | | | (0.020) | | | | | (0.027) |
| $\xi$ | | -0.117 | -0.122 | -0.238 | -0.080 | -0.038 | -0.033 | -0.020 | -0.025 | -0.065 | -0.107 | -0.238 | -0.416 |
| | | (0.040) | (0.050) | (0.041) | (0.045) | (0.083) | (0.046) | (0.050) | (0.039) | (0.041) | (0.050) | (0.054) | (0.046) |
| $\psi$ | | 0.915 | 0.952 | 0.933 | 0.936 | 0.937 | 0.916 | 0.931 | 0.932 | 0.924 | 0.916 | 0.972 | 0.911 |
| | | (0.052) | (0.060) | (0.050) | (0.051) | (0.107) | (0.052) | (0.060) | (0.070) | (0.055) | (0.055) | (0.069) | (0.053) |
| $\sigma_u$ | | 0.343 | 0.342 | 0.409 | 0.341 | 0.302 | 0.300 | 0.304 | 0.309 | 0.340 | 0.412 | 0.497 | 0.340 |
| | | (0.080) | (0.080) | (0.080) | (0.080) | (0.080) | (0.080) | (0.080) | (0.080) | (0.080) | (0.080) | (0.080) | (0.080) |
| $\tau_1$ | | 0.001 | -0.001 | -0.011 | -0.004 | 0.004 | 0.000 | -0.009 | -0.009 | 0.008 | 0.014 | 0.016 | -0.012 |
| | | (0.013) | (0.013) | (0.016) | (0.013) | (0.011) | (0.012) | (0.012) | (0.013) | (0.013) | (0.015) | (0.019) | (0.022) |
| $\tau_2$ | | 0.101 | 0.091 | 0.135 | 0.107 | 0.067 | 0.072 | 0.064 | 0.066 | 0.068 | 0.075 | 0.075 | 0.149 |
| | | (0.009) | (0.010) | (0.011) | (0.009) | (0.011) | (0.008) | (0.008) | (0.008) | (0.009) | (0.010) | (0.013) | (0.016) |
| $l(r,x)$ | | -1241.678 | -1238.017 | -1371.641 | -1236.024 | -1147.060 | -1138.217 | -1137.100 | -1160.411 | -1231.604 | -1379.533 | -1517.672 | -1630.638 |
| $l(r)$ | -993.166 | -980.270 | -979.957 | -979.023 | -979.652 | -980.544 | -977.052 | -976.588 | -977.829 | -976.440 | -980.523 | -978.992 | -971.308 |
| out-of-sample | | | | | | | | | | | | | |
| $\alpha_{MZ}$ | 0.418 | 0.343 | 0.220 | 0.559 | 0.328 | 0.193 | 0.201 | 0.172 | 0.254 | 0.159 | 0.099 | 0.129 | 0.116 |
| $\beta_{MZ}$ | 0.714 | 0.700 | 0.818 | 0.390 | 0.765 | 0.929 | 0.922 | 0.954 | 0.869 | 0.969 | 1.031 | 1.014 | 1.051 |
| $R^2_{MZ}$ | 0.008 | 0.142 | 0.208 | 0.046 | 0.150 | 0.288 | 0.448 | 0.451 | 0.425 | 0.408 | 0.361 | 0.299 | 0.260 |
| HMSE | 3.387 | 2.082 | 1.980 | 1.444 | 3.406 | 1.654 | 1.239 | 1.249 | 1.298 | 1.432 | 1.724 | 1.980 | 2.088 |
| QLIKE | 1.297 | 1.190 | 1.162 | 1.044 | 1.340 | 1.293 | 1.216 | 1.219 | 1.219 | 1.234 | 1.260 | 1.279 | 1.298 |

Table 4: Results for the EUR futures: in-sample fits of GARCH(1,1), Realized GARCH(1,1) with RV, BV, RK, TSRV, JWTSRV, Realized (Wavelet)Jump-GARCH denoted as RJ-G and RWJ-G and Realized Jump-GARCH on $JWTSRV_j$ decompositions. Robust standard errors are reported in parentheses.

| | GARCH | Realized GARCH | | | | | Realized (W)J-GARCH | | Realized Jump-GARCH on $JWTSRV_j$ | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| In-sample | | RV | BV | RK | TSRV | JWTSRV | RJ-G | RWJ-G | $j=1$ | $j=2$ | $j=3$ | $j=4$ | $j=5$ |
| $\omega$ | 0.003 | 0.084 | 0.114 | 0.089 | 0.083 | 0.150 | 0.133 | 0.159 | 0.115 | 0.112 | 0.120 | 0.097 | 0.091 |
| | (0.002) | (0.027) | (0.025) | (0.019) | (0.021) | (0.027) | (0.033) | (0.035) | (0.028) | (0.027) | (0.030) | (0.021) | (0.018) |
| $\alpha$ | 0.055 | | | | | | | | | | | | |
| | (0.012) | | | | | | | | | | | | |
| $\beta$ | 0.945 | 0.676 | 0.628 | 0.744 | 0.673 | 0.508 | 0.511 | 0.512 | 0.535 | 0.573 | 0.657 | 0.752 | 0.813 |
| | (0.010) | (0.038) | (0.040) | (0.032) | (0.037) | (0.044) | (0.044) | (0.042) | (0.041) | (0.045) | (0.036) | (0.027) | (0.024) |
| $\gamma$ | | 0.305 | 0.337 | 0.228 | 0.299 | 0.423 | 0.436 | | | | | | |
| | | (0.042) | (0.042) | (0.029) | (0.038) | (0.046) | (0.045) | | | | | | |
| $\gamma_J$ | | | | | | | 0.132 | 0.133 | 0.137 | 0.151 | 0.105 | 0.106 | 0.095 |
| | | | | | | | (0.062) | (0.067) | (0.054) | (0.062) | (0.055) | (0.069) | (0.073) |
| $\gamma_{W_1}$ | | | | | | | | 0.188 | 0.413 | | | | |
| | | | | | | | | (0.069) | (0.044) | | | | |
| $\gamma_{W_2}$ | | | | | | | | 0.058 | | 0.387 | | | |
| | | | | | | | | (0.062) | | (0.042) | | | |
| $\gamma_{W_3}$ | | | | | | | | 0.101 | | | 0.295 | | |
| | | | | | | | | (0.035) | | | (0.040) | | |
| $\gamma_{W_4}$ | | | | | | | | 0.022 | | | | 0.158 | |
| | | | | | | | | (0.028) | | | | (0.026) | |
| $\gamma_{W_5}$ | | | | | | | | 0.061 | | | | | 0.158 |
| | | | | | | | | (0.023) | | | | | (0.022) |
| $\xi$ | | -0.281 | -0.342 | -0.396 | -0.282 | -0.359 | -0.336 | -0.308 | -0.311 | -0.328 | -0.444 | -0.510 | -0.630 |
| | | (0.061) | (0.055) | (0.052) | (0.053) | (0.048) | (0.059) | (0.062) | (0.049) | (0.053) | (0.060) | (0.065) | (0.056) |
| $\psi$ | | 1.036 | 1.074 | 1.095 | 1.064 | 1.132 | 1.083 | 1.074 | 1.083 | 1.062 | 1.120 | 1.128 | 1.130 |
| | | (0.056) | (0.068) | (0.054) | (0.060) | (0.072) | (0.058) | (0.065) | (0.060) | (0.063) | (0.086) | (0.074) | (0.061) |
| $\sigma_u$ | | 0.330 | 0.326 | 0.391 | 0.335 | 0.296 | 0.291 | 0.289 | 0.294 | 0.322 | 0.369 | 0.451 | 0.322 |
| | | (0.080) | (0.080) | (0.080) | (0.080) | (0.080) | (0.080) | (0.080) | (0.080) | (0.080) | (0.080) | (0.080) | (0.080) |
| $\tau_1$ | | -0.005 | -0.009 | -0.007 | -0.006 | -0.021 | -0.023 | -0.031 | -0.029 | -0.019 | -0.015 | -0.024 | -0.010 |
| | | (0.012) | (0.012) | (0.015) | (0.013) | (0.011) | (0.011) | (0.011) | (0.011) | (0.012) | (0.014) | (0.017) | (0.021) |
| $\tau_2$ | | 0.086 | 0.077 | 0.114 | 0.091 | 0.058 | 0.067 | 0.058 | 0.060 | 0.065 | 0.069 | 0.059 | 0.142 |
| | | (0.009) | (0.008) | (0.011) | (0.009) | (0.008) | (0.008) | (0.008) | (0.007) | (0.008) | (0.009) | (0.011) | (0.015) |
| $l(r,x)$ | | -1189.910 | -1181.787 | -1317.188 | -1201.195 | -1111.835 | -1086.624 | -1079.101 | -1095.175 | -1161.609 | -1261.630 | -1410.358 | -1541.190 |
| $l(r)$ | -967.463 | -958.520 | -958.075 | -958.107 | -957.779 | -960.222 | -947.782 | -947.268 | -948.670 | -947.340 | -946.237 | -944.392 | -946.267 |
| out-of-sample | | | | | | | | | | | | | |
| $\alpha_{MZ}$ | -2.104 | 0.005 | -0.055 | -0.029 | -0.025 | -0.125 | -0.072 | -0.072 | -0.036 | -0.077 | -0.127 | -0.128 | -0.234 |
| $\beta_{MZ}$ | 3.103 | 0.910 | 0.928 | 0.881 | 0.942 | 0.976 | 0.936 | 0.932 | 0.910 | 0.944 | 1.019 | 1.024 | 1.087 |
| $R^2_{MZ}$ | 0.381 | 0.505 | 0.552 | 0.402 | 0.485 | 0.596 | 0.613 | 0.614 | 0.592 | 0.596 | 0.547 | 0.509 | 0.501 |
| HMSE | 3.206 | 1.308 | 1.132 | 1.113 | 1.558 | 1.088 | 0.987 | 0.969 | 1.078 | 1.053 | 1.375 | 1.604 | 1.602 |
| QLIKE | 1.167 | 0.995 | 0.915 | 0.903 | 1.009 | 0.890 | 0.875 | 0.875 | 0.875 | 0.882 | 0.906 | 0.916 | 0.939 |

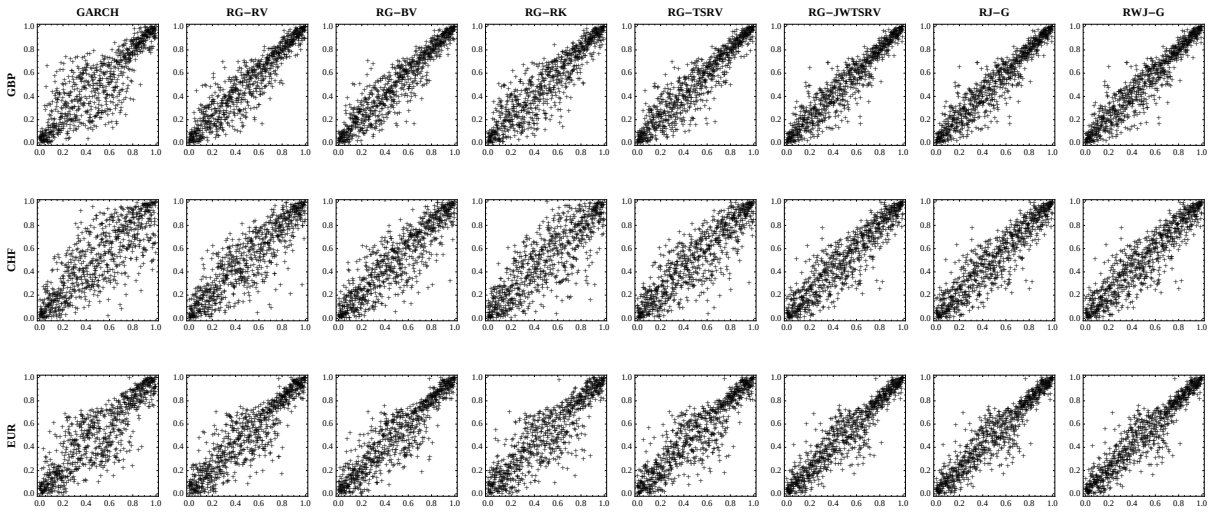


Figure 3: Scattered plot of $h_t$ on $x_t$ mapped into probability integral transform (PITs) for all different models. Rows contain estimates of GBP, CHF and EUR futures, while columns contain GARCH(1,1), Realized GARCH(1,1) estimates using RV, BV, RK, TSRV, JWTSRV denoted as RG-RV, RG-BV, RG-RK, RG-TSRV, RG-JWTSRV, and finally Realized Jump-GARCH(1,1) and Realized Wavelet Jump-GARCH(1,1) denoted as RJ-G and RWJ-G respectively.

Tables 2, 3 and 4 contain all results for GBP futures, CHF futures and EUR futures respectively. By observing partial log-likelihood $\ell(r)$, we can see immediately that all the Realized GARCH models reported by the second, third, fourth, fifth and sixth columns bring significant improvement to the GARCH(1,1) model reported by the first column (in testing significance of the difference, we restrict ourselves to use simple log-likelihood ratio test). When we focus on comparison of Realized GARCH(1,1) with different realized measures $x_t$, we observe further significant differences. This points to importance of usage proper realized measure. While the simplest measure RV is contammined with noise and jumps, we expect the worst performance for the model which uses it as a realized measure proxy. BV is robust to jumps and RK with TSRV are robust to noise. Finally, our JWTSRV estimator is robust to both jumps and noise in the realized variance so we expect the best performance of model which uses JWTSRV. Looking at the results, all the parameter estimates for the different realized measures are similar to each other, while log-likelihoods $\ell(r, x)$ uncover rather large differences between the models. In all three currency futures used in this study, Realized GARCH(1,1) model with JWTSRV realized measure performs significantly better than in RV, BV, RK and TSRV cases. Its log-likelihood brings the largest improvement to all other models. Models with RV, BV and TSRV are more or less on the similar levels of the log-likelihood, while surprisingly the model with RK measure of realized variance is far worst in all cases.

Figure 3 compares the latent volatility $h_t$ and measured volatility $x_t$ from all models. It brings further insight into the various fits and it confirms our findings. When compared to RV, BV, RK and TSRV measures, we can see that relationship between $h_t$ and $x_t$ is strongest for our last three models based on the JWTSRV measure. Moreover the plot for RK explains why it performs so badly in comparison to other estimators. Figure 4 shows the scattered plots of residuals $z_t$ and $u_t$ and confirms a good specification of all models.

Knowing that Realized GARCH(1,1) with the JWTSRV measure performs far best in all cases

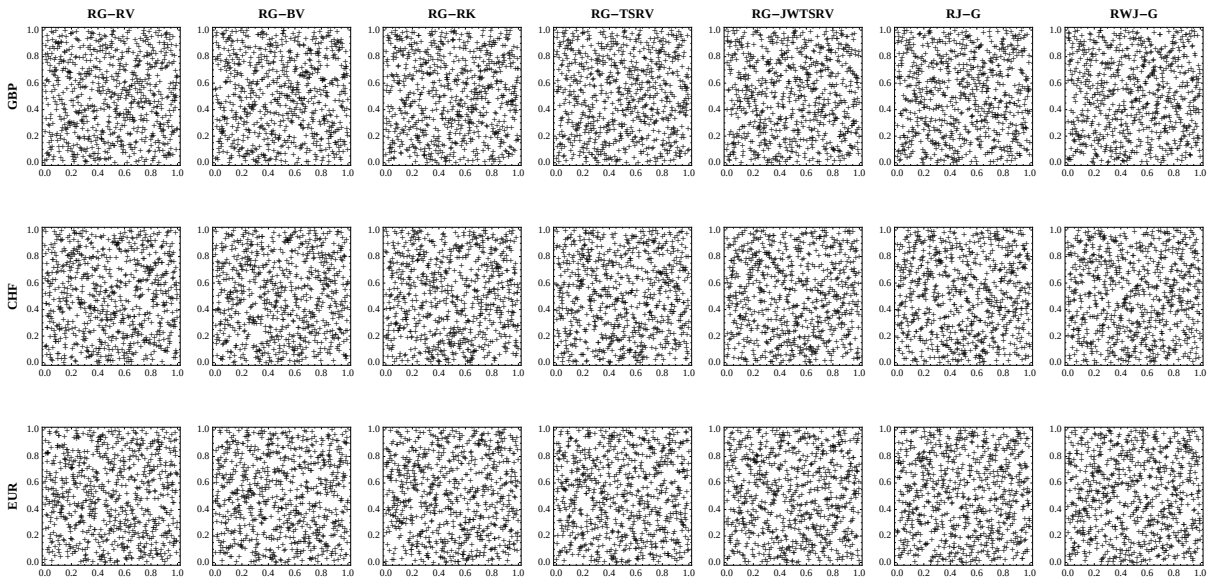


Figure 4: Scattered plot of $z_t$ on $u_t$ residuals mapped into probability integral transform (PITs) obtained for different models. Rows contain estimates of GBP, CHF and EUR futures, while columns contain Realized GARCH(1,1) estimates using RV, BV, RK, TSRV, JWTSRV denoted as RG-RV, RG-BV, RG-RK, RG-TSRV, RG-JWTSRV, and finally Realized Jump-GARCH(1,1) and Realized Wavelet Jump-GARCH(1,1) denoted as RJ-G and RWJ-G respectively.

and improves the log-likelihood supports our further modifications. Motivated by these results, we study if inclusion of jumps in the model improves the fits in our newly proposed Realized Jump-GARCH(1,1) model denoted as Realized J-GARCH in the Tables. Realized J-GARCH model brings further significant improvement in the log-likelihood in all cases, while $\gamma_J$ coefficient is significantly different from zero in the case of CHF and EUR, but can not be statistically distinguished from zero in the case of GBP. The only reason we can see is that in case of GBP futures the estimated jump variation is lowest in comparison to other currencies used so it does not play significant role in forecasts. Still, we can conclude that jumps bring significant improvement in the modeling and Realized Jump-GARCH(1,1) using JWTSRV outperforms other models.

As the last step, we would like to utilize the realized variance decomposition of JWTSRV as we expect that it will further improve the forecasts. Our motivation is straightforward. We would like to find out if the different investment horizons bring improvement to the volatility forecasts. For this, we utilize our newly proposed Realized Wavelet Jump-GARCH(1,1) model (Eq. 27) denoted as Realized WJ-GARCH, where we include all estimated components of realized variation, i.e. all different components of JWTSRV and jumps. Results for all three currencies are strikingly conclusive for the Realized WJ-GARCH model. It brings further improvement in the both full and partial log-likelihoods. Interestingly, $\gamma_{W_1}$ and $\gamma_{W_3}$ coefficients are significantly different from zero in all three cases, while $\gamma_{W_5}$ is significantly different from zero in two cases and $\gamma_{W_2}$ in one case. Significance of jump term does not change. This points us to the conclusion that wavelet decomposition brings not only improvement in overall performance, but it can also be utilized in improvement of the volatility modeling.

The resting five models in the Tables 2, 3 and 4 are Realized Jump-GARCH(1,1) on JWTSRV decompositions separately. Specifically, we use $JWTSRV_j$ for $x_t$ to see the contribution of different decompositions. These last models confirm the previous findings. Interestingly, when only first component (investment horizon of 10 minutes) is used, the model brings very similar performance to the best Realized WJ-GARCH model. This confirms the intuition that most of the information is carried within this scale.

Until now, we have been focusing on in-sample results. Turing our attention to the out-of-sample results we can see that they confirm our findings from the in-sample estimation. Realized WJ-GARCH model improves out-of-sample forecasts in terms of $R^2$ from the Mincer-Zarnowitz regression substantially in comparison to all other models, while HMSE and QLIKE confirms this result. We also note that out-of-sample forecasts are very accurate as $\beta$ from the Mincer-Zarnowitz regressions is very close to 1. It is worth noting that using other realized measures this is not always the case. Still, $\alpha$ shows some bias in the forecasts, especially in case of GBP and CHF currencies. This bias can be contributed to the period we choose for forecasting.

## 6. Conclusion

In this paper, we propose a forecasting model based on decomposed integrated volatilities and jumps. This model utilizes a jump wavelet two scale realized volatility estimator which measures volatility in the time-frequency domain, and recently proposed Realized GARCH model. While the JWTSRV estimator is able to consistently estimate jumps from the price process, it can also be used to decompose the volatility into several investment horizons. This motivates us to propose a new Realized GARCH models including the jumps as well as volatility decomposed into arbitrarily chosen number of investment horizons.

After the introduction of wavelet-based estimation of quadratic variation and all the estimators used in the study, we build a new Realized Jump-GARCH and Realized Wavelet Jump-GARCH models for volatility forecasting. We compare our estimators to several most popular estimators, namely, realized variance, bipower variation, two-scale realized volatility and realized kernels in the forecasting exercise using Realized GARCH framework. Usage of the wavelet-based estimator proves to bring significant improvement in the volatility forecasts. Model incorporating jumps improves forecasting ability even more.

Concluding the empirical findings, we show that our wavelet-based estimators brings a significant improvement to volatility estimation and forecasting. It also offers a new method of time-frequency modeling of realized volatility which helps us to better understand the dynamics of stock market behavior. Specifically, our theory uncovers that most of the volatility is created on higher frequencies.

## References


Andersen, T. and T. Bollerslev (1998). Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *International Economic Review* (39).

Andersen, T., T. Bollerslev, F. Diebold, and P. Labys (2001). The distribution of realized exchange rate volatility. *Journal of the American Statistical Association* (96), 42–55.

Andersen, T., T. Bollerslev, F. Diebold, and P. Labys (2003). Modeling and forecasting realized volatility. *Econometrica* (71), 579–625.

Andersen, T., T. Bollerslev, and X. Huang (2011). A reduced form framework for modeling volatility of speculative prices based on realized variation measures. *Journal of Econometrics 160*(1), 176–189.

Antoniou, I. and K. Gustafson (1999). Wavelets and stochastic processes. *Mathematics and Computers in Simulation 49*, 81–104.

Barndorff-Nielsen, O., P. Hansen, A. Lunde, and N. Shephard (2008). Designing realized kernels to measure the ex-post variation of equity prices in the presence of noise. *Econometrica 76*(6), 1481–1536.

Barndorff-Nielsen, O. and N. Shephard (2001). Non-gaussian ornstein-uhlenbeck-based models and some of their uses in financial economics. *Journal of the Royal Statistical Society, Series B* (63), 167–241.

Barndorff-Nielsen, O. and N. Shephard (2002a). Econometric analysis of realised volatility and its use in estimating stochastic volatility models. *Journal of the Royal Statistical Society, Series B* (64), 253–280.

Barndorff-Nielsen, O. and N. Shephard (2002b). Estimating quadratic variation using realized variance. *Journal of Applied Econometrics* (17), 457–477.

Barndorff-Nielsen, O. and N. Shephard (2004). Power and bipower variation with stochastic volatility and jumps. *Journal of Financial Econometrics* (2), 1–48.

Barndorff-Nielsen, O. and N. Shephard (2006). Econometrics of testing for jumps in financial economics using bipower variation. *Journal of Financial Econometrics* (4), 1–30.

Barunik, J. and L. Vacha (2012). Realized wavelet-based estimation of integrated variance and jumps in the presence of noise. *arXiv preprint* (1202.1854).

Bollerslev, T., R. Engle, and D. Nelson (1994). *Handbook of Econometrics*, Volume IV, Chapter ARCH models, pp. 2961–3038. Elsevier Science B.V.

Bollerslev, T. and E. Ghysels (1996). Periodic autoregressive conditional heteroscedasticity. *Journal of Business & Economic Statistics 14*(2), 139–151.

Capobianco, E. (2004). Multiscale stochastic dynamics in finance. *Physica A* (344), 122–127.

Dacorogna, M., U. Müller, R. Nagler, R. Olsen, and O. Pictet (2001). *An Introduction to High-Frequency Finance*. Academic Press, San Diego.

Daubechies, I. (1992). *Ten lectures on wavelets*. SIAM.

Donoho, D. L. and I. M. Johnstone (1994). Ideal spatial adaptation by wavelet shrinkage. *Biometrica* (81), 425–455.

Fan, J. and Y. Wang (2007). Multi-scale jump and volatility analysis for high-frequency financial data. *Journal of the American Statistical Association* (102), 1349–1362.

Gençay, R., F. Selçuk, and B. Whitcher (2002). *An Introduction to Wavelets and Other Filtering Methods in Finance and Economics*. Academic Press.

Gençay, R., N. Gradojevic, F. Selçuk, and B. Whitcher (2010). Asymmetry of information flow between volatilities across time scales. *Quantitative Finance*, 1–21.

Hansen, P. R., Z. Huang, and H. H. Shek (2011). Realized garch: a joint model for returns and realized measures of volatility. *Journal of Applied Econometrics*.

Hø g, E. and A. Lunde (2003). Wavelet estimation of integrated volatility. *Working Paper. Aarhus School of Business.*.

Lanne, M. (2007). Forecasting realized exchange rate volatility by decomposition. *International Journal of Forecasting* (23), 307–320.

Mallat, S. (1998). *A wavelet tour of signal processing*. Academic Press.

Mancino, M. and S. Sanfelici (2008). Robustness of fourier estimator of integrated volatility in the presence of microstructure noise. *Computational Statistics & data analysis 52*(6), 2966–2989.

Mincer, J. and V. Zarnowitz (1969). *The evaluation of economic forecasts*. New York: National Bureau of Economic Research.

Nielsen, M. and P. Frederiksen (2008). Finite sample accuracy and choice of sampling frequency in integrated volatility estimation. *Journal of Empirical Finance* (15), 265–286.

Olhede, S., A. Sykulski, and G. Pavliotis (2009). Frequency domain estimation of integrated volatility for ito processes in the presence of market-microstructure noise. *Multiscale Modeling & Simulation 8*(2), 393–427.

Percival, D. and A. Walden (1993). *Spectral analysis for physical applications*. Cambridge University Press.

Percival, D. B. (1995). On estimation of the wavelet variance. *Biometrika 82*, 619–631.

Percival, D. B. and H. Mofjeld (1997). Analysis of subtidal coastal sea level fluctuations using wavelets. *Journal of the American Statistical Association 92*(439), 886–880.

Percival, D. B. and A. T. Walden (2000). *Wavelet Methods for Time series Analysis*. Cambridge University Press.

Raimondo, M. (1998). Minimax estimation of sharp change points. *The Annals of Statistics* (26), 1379–1397.

Serroukh, A., A. T. Walden, and D. B. Percival (2000). Statistical properties and uses of the wavelet variance estimator for the scale analysis of time series. *Journal of the American Statistical Association 95*, 184–196.

Subbotin, A. (2008). A multi-horizon scale for volatility. Technical report, Documents de travail du Centre d'Economie de la Sorbonne, Université Panthéon-Sorbonne (Paris 1).

Wang, Y. (1995). Jump and sharp cusp detection via wavelets. *Biometrika* (82), 385–397.

Zhang, L., P. Mykland, and Y. Aït-Sahalia (2005). A tale of two time scales: Determining integrated volatility with noisy high frequency data. *Journal of the American Statistical Association* (100), 1394–1411.

Zhou, B. (1996). High-frequency data and volatility in foreign-exchange rates. *Journal of Business and Economic Statistics 14*, 45–52.

## Appendix A. The maximal overlap discrete wavelet transform

The maximal overlap discrete wavelet transformation (MODWT) is a translation-invariant type of discrete wavelet transformation, i.e., it is not sensitive to the choice of starting point of the examined process. Furthermore, the MODWT does not use a downsampling procedure as in the case of the discrete wavelet transform[4] (DWT), so the wavelet and scaling coefficient vectors at all levels (scales) have equal length. As a consequence, the MODWT is not restricted to sample sizes that are powers of two. This feature is very important for the analysis of real market data, since this limitation is usually too restrictive. In the literature the MODWT is also called the stationary wavelet transform, the translation invariant transform and the undecimated wavelet transform. For more details about the MODWT see Mallat (1998), Percival and Walden (2000) and Gençay et al. (2002).

The MODWT is a very convenient tool for variance and energy analysis of a time series in the time-frequency domain. Percival (1995) demonstrates the advantages of the MODWT estimator of variance over the DWT estimator, Serroukh et al. (2000) analyze the statistical properties of the MODWT variance estimator for non-stationary and non-Gaussian processes.

### *Appendix A.1. Definition of MODWT filters*

First, let us introduce the MODWT scaling and wavelet filters $g_l$ and $h_l$, $l = 0, 1, \ldots, L-1$, where $L$ denotes the length of the wavelet filter. For example, the Daubechies D(4) wavelet filter has length $L = 4$ (Daubechies, 1992). Generally, the scaling filter is a low-pass filter whereas the wavelet filter is a high-pass filter. There are three basic properties that both the MODWT filters must fulfill. Let us show these properties for the MODWT wavelet filter:

$$\sum_{l=0}^{L-1} h_l = 0, \quad \sum_{l=0}^{L-1} h_l^2 = 1/2, \quad \sum_{l=-\infty}^{\infty} h_l h_{l+2N} = 0, \quad N \in \mathbb{Z}_N, \tag{A.1}$$

and for the MODWT scaling filter:

$$\sum_{l=0}^{L-1} g_l = 1, \quad \sum_{l=0}^{L-1} g_l^2 = 1/2, \quad \sum_{l=-\infty}^{\infty} g_l g_{l+2N} = 0, \quad N \in \mathbb{Z}_N. \tag{A.2}$$

The transfer function of a MODWT filter $\{h_l\}$ at frequency $f$ is defined via the Fourier transform as:

$$H(f) \equiv \sum_{l=-\infty}^{\infty} h_l e^{-i2\pi fl} = \sum_{l=0}^{L-1} h_l e^{-i2\pi fl}, \tag{A.3}$$

with the squared gain function defined as: $\mathcal{H}(f) \equiv |H(f)|^2$.

### *Appendix A.2. Pyramid algorithm*

We get the MODWT wavelet and scaling coefficients using the pyramid algorithm. The wavelet coefficients at the first scale ($j = 1$) are obtained via filtering $x_i$ on $i = 1, \ldots, N$ with the MODWT

[4]For a definition and detailed discussion of the discrete wavelet transform see Mallat (1998), Percival and Walden (2000) and Gençay et al. (2002)

wavelet and scaling filters (Percival and Walden, 2000):

$$W_{1,k} \equiv \sum_{l=0}^{L-1} h_l x_{k-lmodN}, \qquad V_{1,k} \equiv \sum_{l=0}^{L-1} g_l x_{k-lmodN}. \tag{A.4}$$

For the second stage of the algorithm, we replace $x_t$ with the scaling coefficients $V_{1,k}$ and after the filtering we get wavelet coefficients at scale $j = 2$ as:

$$W_{2,k} \equiv \sum_{l=0}^{L-1} h_l V_{1,k-lmodN}, \qquad V_{2,k} \equiv \sum_{l=0}^{L-1} g_l V_{1,k-lmodN}. \tag{A.5}$$

After two stages of the pyramid algorithm we have two vectors of the MODWT wavelet coefficients $\mathbf{W}_1, \mathbf{W}_2$ and one vector of the MODWT wavelet scaling coefficients at scale two $\mathbf{V}_2$. Vector $\mathbf{W}_1$ represents wavelet coefficients at the frequency band $f \in [1/4, 1/2]$, $\mathbf{W}_2$: $f \in [1/8, 1/4]$ and $\mathbf{V}_2$: $f \in [0, 1/8]$. The $j$-th level MODWT coefficients are in the form:

$$W_{j,k} \equiv \sum_{l=0}^{L-1} h_l V_{j-1,k-lmodN}, \qquad V_{j,k} \equiv \sum_{l=0}^{L-1} g_l V_{j-1,k-lmodN}, \quad j = 1, 2, \ldots, J^m. \tag{A.6}$$

where $J^m \leq log_2(N)$. Generally, the $j$-th level wavelet coefficients in the vector $\mathbf{W}_j$ represents frequency bands $f \in [1/2^{j+1}, 1/2^j]$ wheres the $j$-th level scaling coefficients in the vector $\mathbf{V}_j$ represents $f \in [0, 1/2^{j+1}]$. For estimation of the wavelet covariance we use MODWT wavelet coefficients unaffected by the boundary conditions. For simplicity in notation let us we define a vector $\mathcal{W}$ that consists of $J^m + 1$ $N-$dimensional subvectors, where the first $J^m$ subvectors are the MODWT wavelet coefficients at levels $j = 1, ..., J^m$ and the last subvector consists of the MODWT scaling coefficients at a level $J^m$:

$$\mathcal{W} = [\mathbf{W}_1, \mathbf{W}_2, \ldots, \mathbf{W}_{J^m}, \mathbf{V}_{J^m}]^T. \tag{A.7}$$

*Appendix A.3. Wavelet decomposition of a stochastic process*

For our analysis, it is important to show that we are able to decompose the variance (energy) of a stochastic process on a scale-by-scale basis, i.e., we can get the variance contribution of every level $j$, with the maximum level of decomposition $J^m \leq \log_2 N$. The (total) variance of the time series $x_i$, $i = 1, \ldots, N$ can be decomposed on a scale-by-scale basis so that

$$\|\mathbf{x}\|^2 = \sum_{j=1}^{J^m} \|\mathbf{W}_j\|^2 + \|\mathbf{V}_{J^m}\|^2 = \sum_{j=1}^{J^m+1} \|\mathcal{W}_j\|^2. \tag{A.8}$$

where $\|\mathbf{x}\|^2 = \sum_{i=1}^{N} x_i^2$, $\|\mathbf{W}_j\|^2 = \sum_{i=1}^{N} W_{j,i}^2$, $\|\mathbf{V}_{J^m}\|^2 = \sum_{i=1}^{N} V_{J^m,i}^2$ and $\mathbf{W}_j$ and $\mathbf{V}_j$ are $N$ dimensional vectors of the $j$-th level MODWT wavelet and scaling coefficients. The proof of the variance decomposition, Eq. (A.8), using the MODWT can be found in Percival and Mofjeld (1997) and Percival and Walden (2000).

It is worth noting that the squared norm $\|.\|$ is similar to the realized measure discussed in the preceding sections. For example, in the case of the realized variance estimator (RV) the energy

decomposition can reveal the contributions of particular scales to the overall energy, hence we can see what form this realized measure takes.